Storage Solutions

Shane Canon , NCCS
March 28, 2007

Outline

- **Overview of File System at LCF**
  - Home File System (Spin/NetApp)
  - Local Scratch file system (XT3, Ewok, *etc*)
  - Spider File System

- **Archival Storage – HPSS**

- **Data Tools**

- **Future Plans**
  - Spider
  - HPSS

Home File Server

- **Purpose: Home directories and project space (code, small input data, scripts)**

- **Mounted on all NCCS systems**

- **Quotas**
  - 500 MB for home directories
  - 5 GB defaults for project directories

- **10-20 MB/s single stream bandwidth**

- **Backed up and snapshots**

- **Future:**
  - Will begin transitioning to a NetApp based system over the next few months
  - New system should be roughly 3x – 4x faster (metadata and block transfers)
  - Quotas boosted to 2 GB

Local File Systems - Jaguar

- **Purpose: High-performance scratch file system**

- **Configuration:**
  - Lustre
  - 3 separate file systems (two 150 TB file systems, one 300 TB file system)
  - Measured 26 GB/s Write and 43 GB/s Read to a single shared file on the 300 TB file system

- **Policies**
  - Files over 7 days are eligible to be purged

- **Not Backed up**
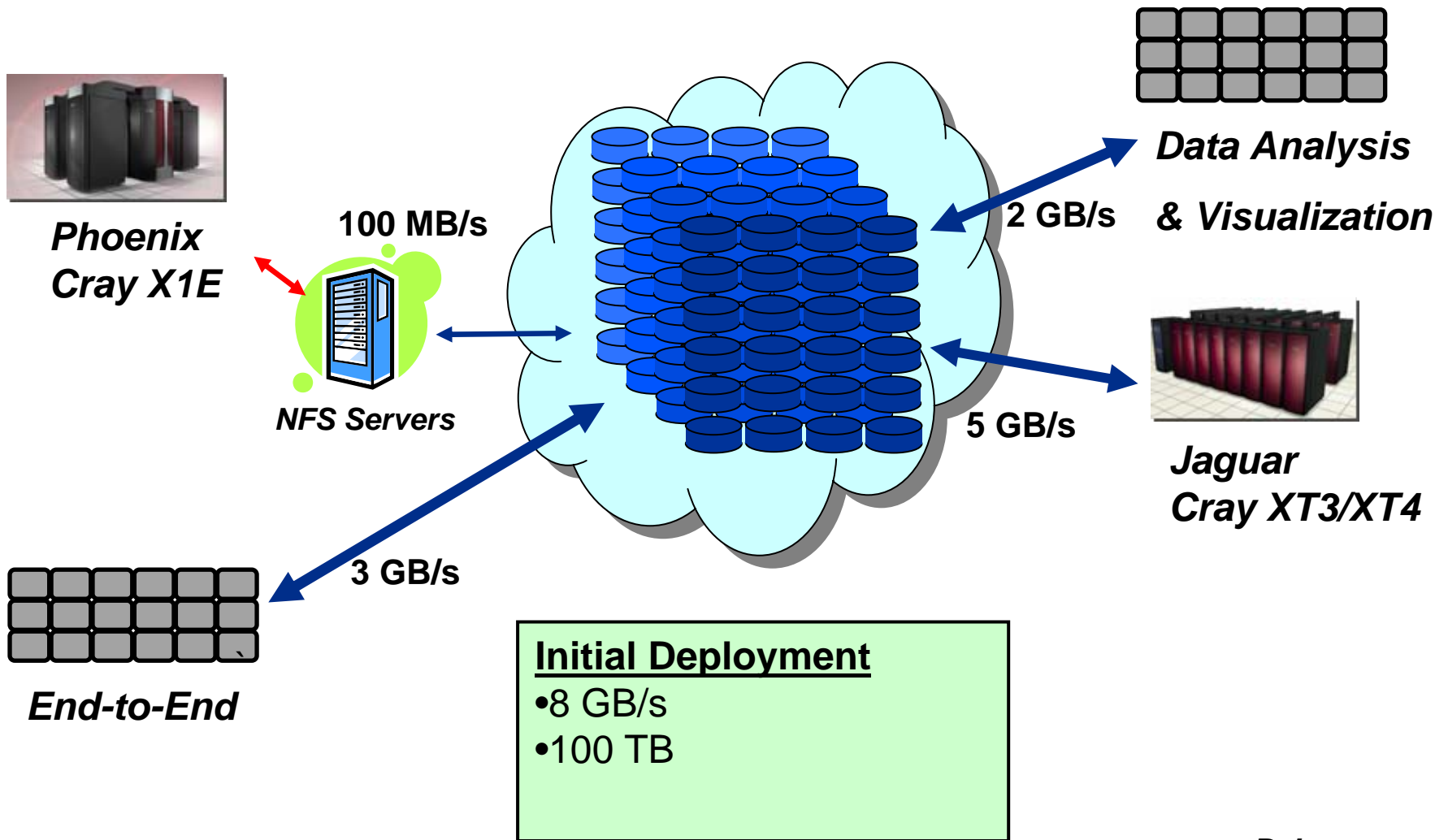
Other Local Scratch File Systems

- **Local scratch file systems also on…**
  - Phoenix (XFS)
  - Ewok (Lustre and local disk)
  - Ram (XFS)
  - Hawk (NFS and local disk)

Spider

- **Purpose: High Performance Center-wide File System**

- **Configuration**
  - Lustre based

- **Policies**
  - Quotas – Specific amounts not determined yet, but will likely be divided over projects

- **Connected to Platforms via 10Gb and IB**

- **Performance**
  - Initial system will provide around 8-10 GB/s of aggregate bandwidth

Spider – Center-Wide File System

Phoenix Cray X1E

100 MB/s

NFS Servers

Data Analysis & Visualization

2 GB/s

5 GB/s

Jaguar Cray XT3/XT4

3 GB/s

End-to-End

Initial Deployment
- 8 GB/s
- 100 TB

Baker

Archival Storage - HPSS

- **Purpose: Long-term stable storage (but not WO-RN)**

- **Configuration**
  - 4 Tape Silos (soon 6)
  - Latest drives store around 500 GB compressed

- **Policies**
  - Currently unconstrained…
  - Plan to implement allocations in near future

- **Connected to 10Gb Ethernet network**

- **Performance**
  - Single file transfer rates over 500 MB/s to/from disk cache
  - Small file performance is poor…please use tar/htar to aggregate smaller files prior to storing in HPSS (think archive not file system)

Summary of Storage Solutions

File System	Purged	Backed Up	Quotas	Global	Purpose
/home	N	Y	Y	Y	Code, scripts
/tmp/work	Y (7 days)	N	N	N	Check-points
Spider	N	N	Likely	Y	
HPSS	N	(It is the backup)	N(ot yet)	Y	Archive

Data Tools

- **Parallel Copy tools**
  - Efficiently copy whole directory trees from one Lustre file system to another
  - Use multiple clients to sustain high transfer rates

- **IOTA – IO Toolkit and Analyzer**
  - Currently supports "hints" file to control striping
  - Plan to add some analysis functions to suggest methods of improving performance based on observed IO patterns
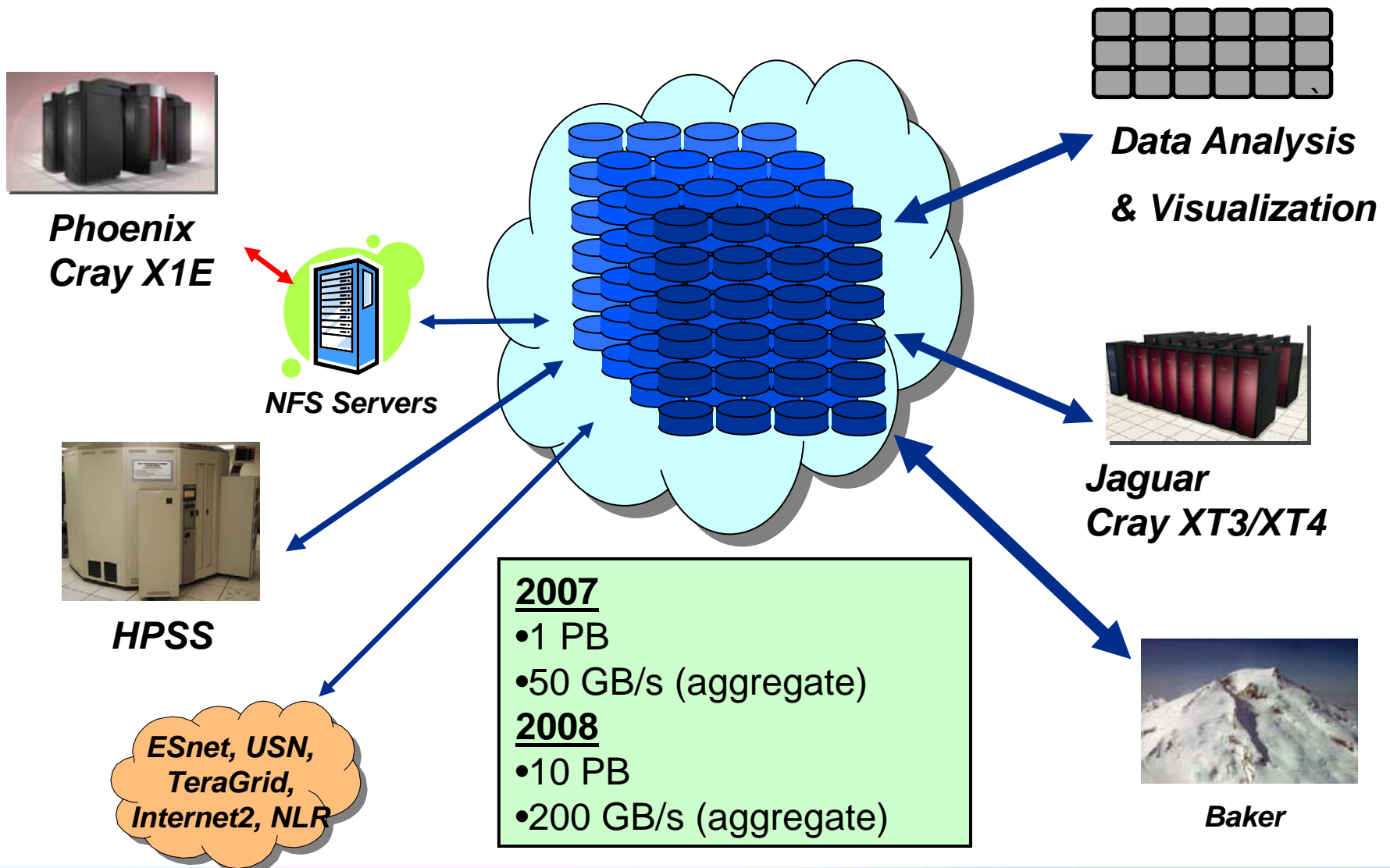
- **MPI-IO**
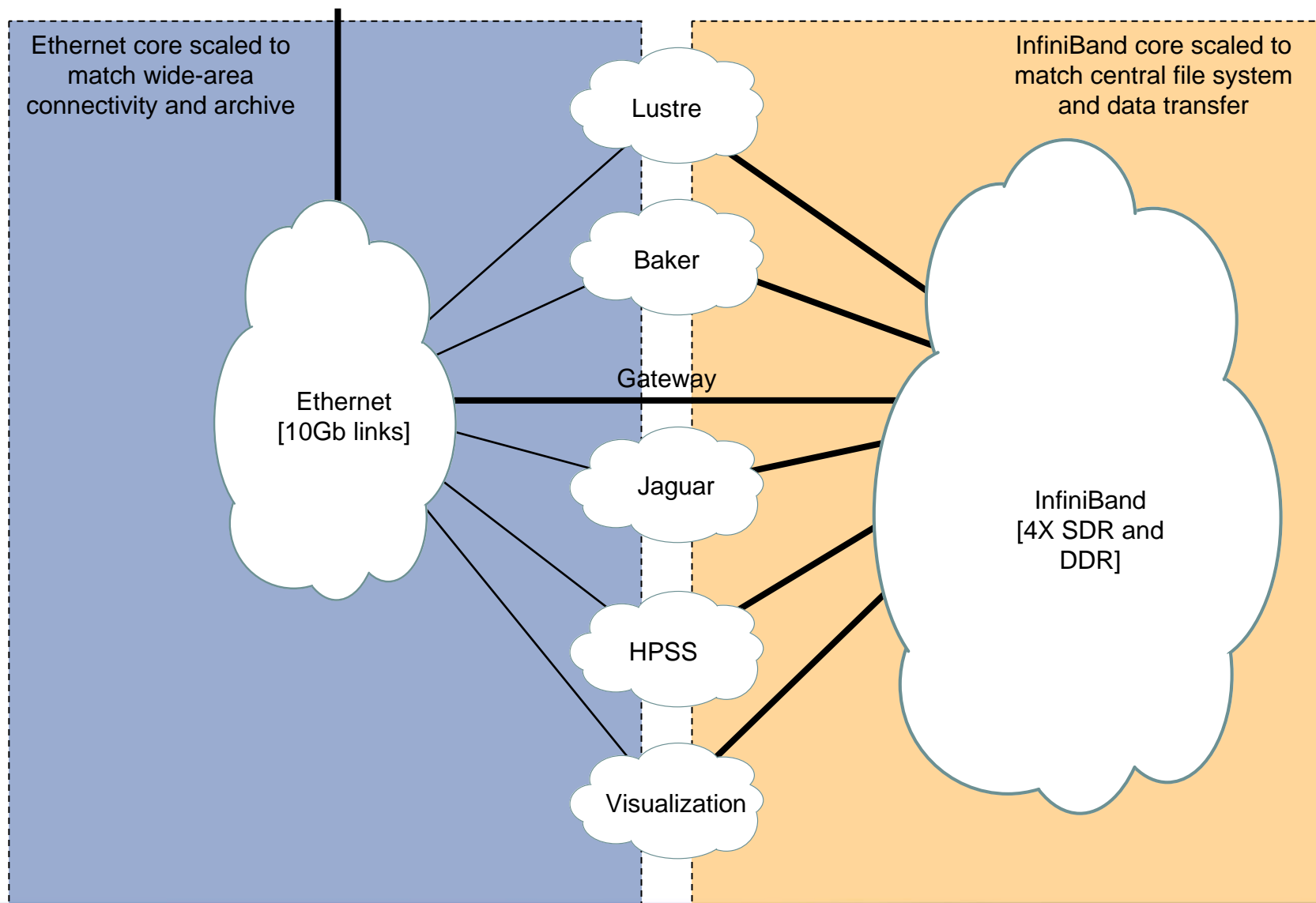 - Targeted Optimization of some MPI-IO collectives

- **Grid**
  - Tested integration of RSA fobs with my-proxy server

Future Plans for Spider

Phoenix Cray X1E

NFS Servers

HPSS

ESnet, USN, TeraGrid, Internet2, NLR

Data Analysis & Visualization

Jaguar Cray XT3/XT4

Baker

2007
- 1 PB
- 50 GB/s (aggregate)

2008
- 10 PB
- 200 GB/s (aggregate)

NCCS Network Roadmap Summary

Ethernet core scaled to match wide-area connectivity and archive

InfiniBand core scaled to match central file system and data transfer

Lustre

Baker

Gateway

Ethernet [10Gb links]

Jaguar

InfiniBand [4X SDR and DDR]

HPSS

Visualization

Spider/HPSS Integration

- **Exploring new techniques to improve data movement between the file system and archival storage**

- **New techniques will take advantage of parallelism in both Lustre and HPSS**

- **Also working with other Lustre sites to develop an HSM that would integrate with HPSS**

Questions for you

Productivity

- How do you spend your time working on LCF related projects? (New Science, Data Management, Debugging, Tuning, Scaling, etc)

- Where do you feel you spend too much time?

Center Needs

- What are your needs for
  - external network connectivity
  - disk storage (capacity and bandwidth)
  - archival storage
  - data analysis and viz
  - tools

Other Questions for you

TOOLS

What are you looking for in a development environment?

- code assistance (such as context sensitive help on prototype)
- refactoring (such as rename, move, extract methods etc.)
- integrated debugging
- scalable debugging (what scale)
- static code analysis
- code transformation and optimization
- performance analysis

Are there tools you already use (and love)?

Questions for Me?